URBAN LOCAL CLIMATE ZONE CLASSIFICATION WITH A RESIDUAL CONVOLUTIONAL NEURAL NETWORK AND MULTI-SEASONAL SENTINEL-2 IMAGES

C. P. Qiu[a], M. Schmitt[a], L.C. Mou[a], X. X. Zhu[a,b]

[a] Signal Processing in Earth Observation, Technical University of Munich (TUM), Arcisstr. 21, 80333 Munich, Germany - (chunping.qiu, m.schmitt)@tum.de

[b] Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Oberpfaffenhofen, 82234 Wessling, Germany - xiao.zhu@dlr.de

KEY WORDS: Local Climate Zones (LCZs), Sentinel-2, Spectral features, Classification, Multi-seasonal, Residual Convolutional Neural Network (ResNet)

ABSTRACT:

This study presents a classification framework for the urban Local Climate Zones (LCZs) based on a Residual Convolutional Neural Network (ResNet) architecture. In order to make full use of the temporal and spectral information contained in modern Earth observation data, multi-seasonal Sentinel-2 images are exploited. After training the ResNet, independent predictions are made from the multi-seasonal images. Subsequently, the seasonal predictions are fused in a decision fusion step based on majority voting. A systematical experiment is carried out in a large scale study area located in the center of Europe. A significant accuracy improvement can be achieved by applying majority voting on multi-seasonal predictions. Based on the results, the main challenges and possible solutions of urban LCZ classification are further discussed, providing guidance for large scale urban LCZ mapping.

1. INTRODUCTION

Local Climate Zone (LCZ) mapping (Stewart and Oke, 2012), originally developed for meta-data communication of observational Urban Heat Island (UHI) studies, has gained great interest in the field of remote sensing. The 17 LCZ classes, as displayed in Fig. 1, are based on climate-relevant surface properties mainly related to 3D surface structure (e.g. height and density of buildings and trees) as well as surface cover (e.g., pervious or impervious). Recently, researchers have started to use the LCZ scheme to classify the internal structure of urban areas, providing critical support for various applications such as urban climatology, infrastructure planning, disaster mitigation, etc.

Supervised classification methods using remote sensing data as input provide a valuable tool for LCZ mapping and have been widely studied in the 2017 Data Fusion Contest organized by the Image Analysis and Data Fusion Technical Committee (IADFTC) of the IEEE Geoscience and Remote Sensing Society (GRSS) (Yokoya et al., 2017), among which deep learning methods also played a significant role (Xu et al., 2017). As with image recognition, object detection and semantic segmentation, LCZ classification can benefit from the learned mid- and high-level abstract features from the original images (Zhu et al., 2017).

Aiming for global LCZ mapping for which training data is costly and resource intensive to collect, our work intends to provide an answer to these questions: What LCZ classification accuracy can be achieved if evaluation is carried out in a cross validation manner, i.e., what is the generalization potential of the classifier? For simplicity, we focus on the first ten LCZ classes, which are referred to as urban LCZ classes in this paper. In addition, we add two background classes, namely vegetation and water, to achieve land cover completeness. As a framework for our investigations, we use a residual convolutional neural network (ResNet) (He et al., 2016). The features fed to the ResNet at training time are globally available Sentinel-2 images prepared in a multi-seasonal manner (Radoux et al., 2016). The results achieved in this paper also provide perspectives for other large scale classification tasks.

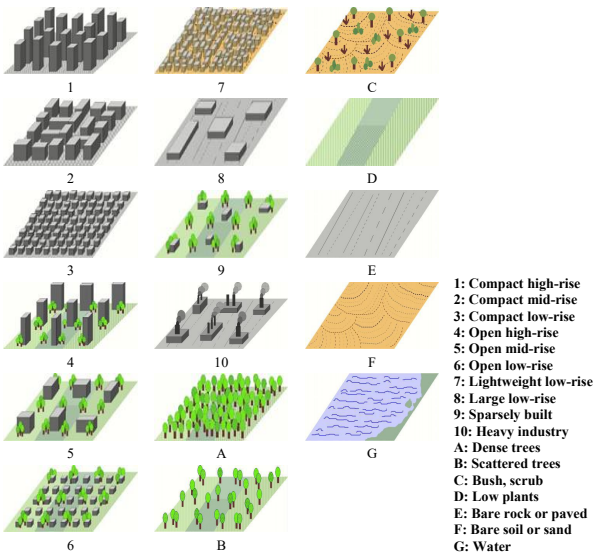


Figure 1. Visualization of the LCZ concept (Stewart, 2011).

2. RESNET FRAMEWORK FOR MULTI-SEASONAL LCZ CLASSIFICATION

We make use of a ResNet, instead of conventional convolutional network, for spatial-spectral classification of LCZs, as ResNets have been proven to be capable of offering better classification performance (Mou et al., 2017). The exact architecture of the

ResNet we train is shown in Fig. 2. Overall, it has three residual blocks, and each of them consists of three convolutional layers and a shortcut connection that by-passes two stacked convolutional layers by performing identity mapping, which are then added together with the output of stacked convolutions. We utilize convolutional layers with a very small receptive field of 3×3, and the number of feature maps increases towards deeper blocks, doubling after each block. Max-pooling is performed over 2×2 pixel windows with stride 2. For training the network, we use the TensorFlow framework. We choose Nesterov Adam as optimization algorithm for our task, as it shows faster convergence than standard stochastic gradient descent with momentum. We fix parameters of Nesterov Adam as recommended: $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 1\text{e}{-08}$, and a schedule decay of 0.004. We use a fairly small learning rate of 0.0002.

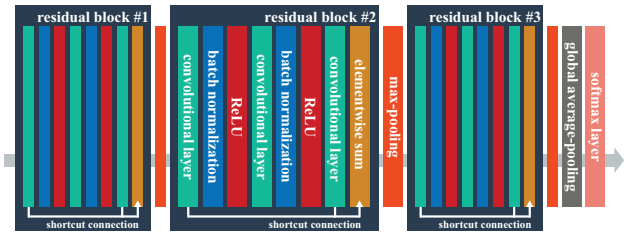


Figure 2. The architecture of the used ResNet for LCZ classification.

Furthermore, we define 15 features as input to the classifier:

1. *Spectral reflectance*
   10 bands of Sentinel-2 imagery are used in this study: B2, B3, B4 and B8 with 10 m Ground Sampling Distance (GSD) and B5, B6, B7, B8a, B11 and B12 with 20 m GSD. The 20 m bands are up-sampled to 10 m GSD. The bands B1, B9 and B10 are not considered in this study because they contain mostly information about the atmosphere and thus bear little relevance to LCZ classification.

2. *Indices*
   The well-established indices Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), Modified Normalized Difference Water Index (MNDWI), Normalized Difference Built Index (NDBI) and Bare-Soil Index (BSI) are also considered (Tucker, 1979), since they can provide indications about vegetation, water, buildings and soil, respectively (Yokoya et al., 2017, Bechtel et al., 2015).

Based on the framework described above, different predictions are retrieved for images acquired in different seasons. On these multi-seasonal predictions a majority voting can be applied to get the final prediction in a decision fusion manner.

3. EXPERIMENTS AND RESULTS

In this section, we describe the test dataset and the achieved experimental results.

3.1 Study Areas and LCZ Dataset

Our study areas are spread over seven cities located in the heart of Europe. They are depicted in Fig. 3. For all those cities, we have downloaded Sentinel-2 imagery from ESA's SciHub (`https://scihub.copernicus.eu/`). One image is used for each city, and altogether there are 28 images.

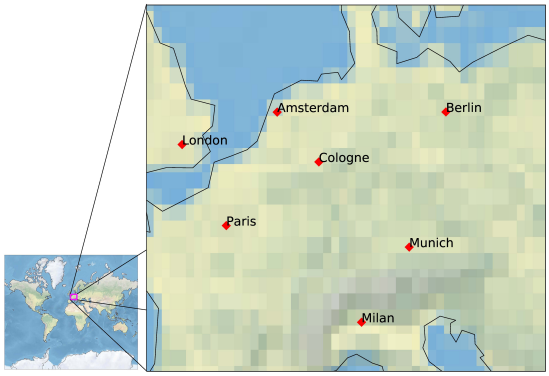


Figure 3. The seven test cities distributed across Europe.

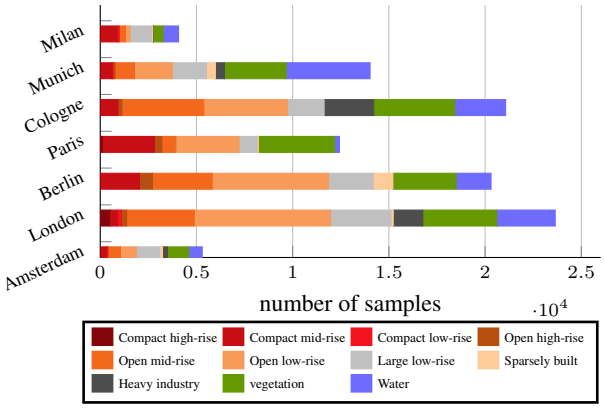


Figure 4. Urban LCZ training sample number of seven cities.

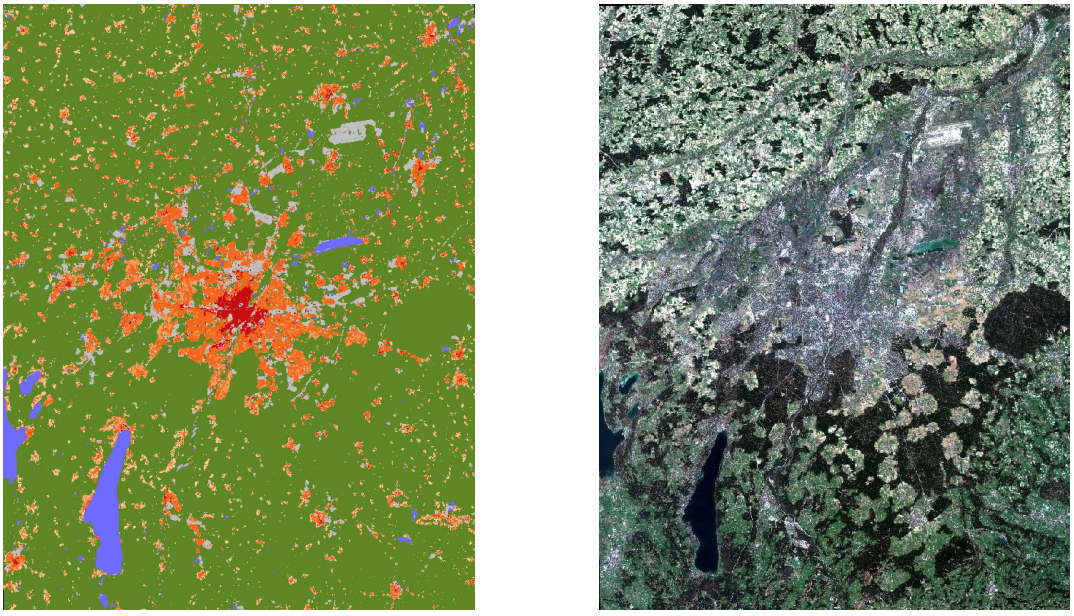
The LCZ ground truth labels available for selected neighborhoods in the seven cities are taken from the LCZ42 dataset (Zhu, 2018). The number of samples available for each class over the seven cities can be seen in Fig. 4, in which the vegetation class combines the LCZ classes A, B, C, D, and F. Figure 4 illustrates the variability of both the sample number and the class distribution among different cities.It should be noted that in these seven cities, LCZ class 7 (lightweight low-rise), which mostly indicates slums, does not exist.

3.2 Classification Results

The classification result from cross validation is shown in Table. 1, where both the majority voting-based result with multi-seasonal Sentinel-2 images (left) and the average of the four single-season results (right) are shown. The achieved classification accuracy is comparable to our previous work, where the same cross validation has been carried out with single-seasonal Sentinel-2 images and additional features (Global Urban Footprint and Open Street Map) (Qiu et al., 2018), using Canonical Correlation Forests as the classifier. Using the majority voting-based result, an exemplary LCZ map over the city of Munich, Germany, is shown in Fig. 5.

Table 1. Classification accuracy of multi-seasonal majority voting and averaged results over four seasons (OA is Overall Accuracy and AA is Average Accuracy).

city	multi-seasonal majority voting			averaged over four seasons		
	OA	Kappa	AA	OA	Kappa	AA
Amsterdam	0.74	0.69	0.53	0.69	0.63	0.49
London	0.66	0.60	0.46	0.63	0.56	0.43
Berlin	0.70	0.64	0.64	0.62	0.55	0.56
Paris	0.82	0.77	0.63	0.74	0.67	0.54
Cologne	0.70	0.64	0.62	0.64	0.57	0.56
Munich	0.76	0.71	0.51	0.76	0.71	0.51
Milan	0.83	0.79	0.58	0.71	0.64	0.50
Mean	**0.74**	**0.69**	**0.57**	**0.68**	**0.62**	**0.51**



(a) (b)

Figure 5. The LCZ map of Munich (a) produced using majority voting, and the corresponding Sentinel-2 RGB imagery (b). The legend is the same with that in Fig. 4

4. DISCUSSION

4.1 Benefits of multi-seasonal classification

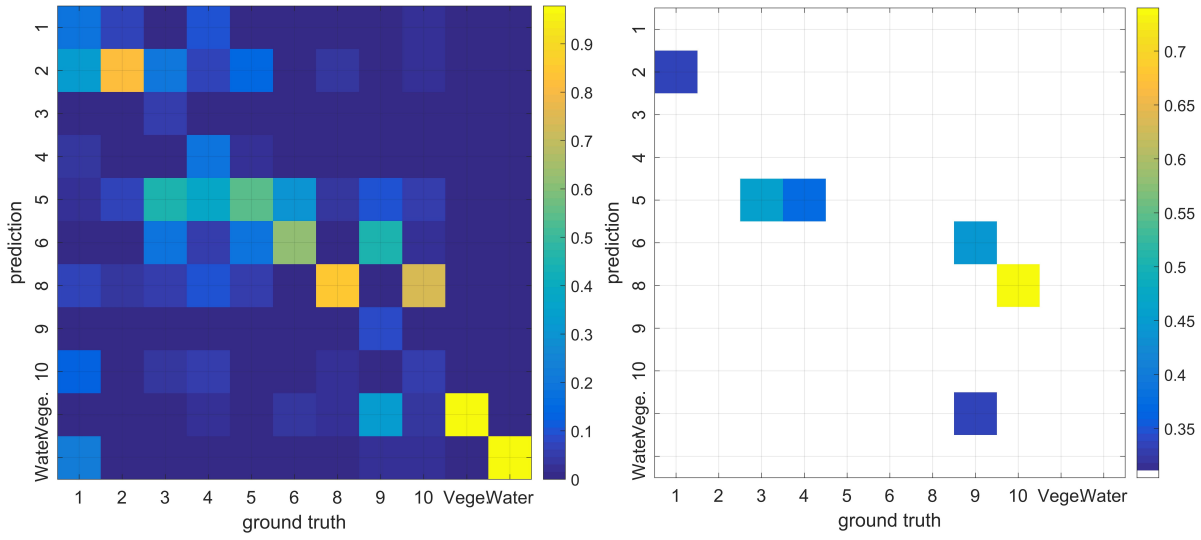
The result achieved in Section 3.2 shows that a promising accuracy can be achieved by applying majority voting on predictions from multi-seasonal images. As the values in Table. 1 show, about 6 percentage points of improvement of overall accuracy, Kappa coefficient and averaged accuracy can be achieved in this case. This improvement, on the one hand, results from the joint distinction capability provided by the temporal and spectral information. On the other hand, this is a piece of evidence for "the wisdom of the crowd": majority voting removes potential outliers in the predictions.

Furthermore, it has to be highlighted that this multi-seasonal result relies only on Sentinel-2 data, and only on four independent, multi-seasonal predictions, which are fused in a simple decision fusion manner based on majority voting. This is in contrast to the more complex set-up that has been necessary in our previous work (Qiu et al., 2018), for which the majority voting-based fusion of 20 different CCF classifier results and external auxiliary data (the Global Urban Footprint and two Open Street Map layers) were necessary. Especially since these Open Street Map layers are not available globally, this shows the great potential of the multi-seasonal ResNet-based approach presented in this paper for global applicability and generalization.

4.2 Misclassification among LCZs

Despite of the promising accuracy of multi-seasonal classification, the average accuracy is still less than 60% for about four of the seven test cities. The misclassification among classes can be analyzed using confusion matrices of the classification results, as shown in Figure 6.

Figure 6(a) depicts the combined confusion matrix of all the seven test cases, and Figure 6(b) highlights the misclassification errors higher than 30%. Classes 1 (compact high rise), 3 (compact low rise) and 4 (open high rise) are falsely classified into classes 2 (compact middle rise), 5 (open middle rise) and 5, respectively. This is resulted from the challenge of distinguishing height differences with Sentinel-2 images. For example, high rise and middle rise are of quite similar appearance in two dimensional optical images. Besides, class 9 (sparsely built) is falsely classified into classes 6 (open low rise) and vegetation, while class 10 (heavy



(a) (b)

Figure 6. Combined confusion matrix of seven cities (a) and the cases with a misclassification error higher than 30% (b).

industry) is falsely classified into class 8 (large low rise). This is due to inter-class similarity, as they appear very similar, as can be seen in Fig. 1.

Another possible reason for the misclassification is that the number of samples is much lower for some classes than for other classes. This class imbalance can be seen in Fig. 4. As a result, the intra-class variability is not well learned during training. These findings are similar to experiences described in previous literature (Yokoya et al., 2017). Figure 7 illustrates the relation between the class sample number and the omission error of each class.It indicates that the classes with fewer samples end up with a higher omission error, which is also confirmed by the relatively large correlation coefficient of -0.75. For example, class 1 (compact high rise), 3 (compact low rise), 4 (open high rise), 9 (sparsely built) and 10 (heavy industry) which are mentioned above, are exactly the classes with higher omission error and few samples.

In order to solve these problems, one possible solution is to include additional datasets such as Synthetic Aperture Radar images to make use of radar's unique range measurements, providing powerful ability to differentiate height. Another solution is to adapt the LCZ scheme considering the feasibility of optical images. Last but not least, negative human influence on ground truth should be mitigated to guarantee the quality of the training samples across cities (Bechtel et al., 2017).

5. SUMMARY AND CONCLUSION

This paper presents an investigation of a ResNet-based multi-seasonal classification framework for urban LCZs with Sentinel-2 images. It intends to evaluate the potential of the generalization ability of the ResNet classifier. A series of experiments has been carried out over seven cities in Europe. The experimental results show that the a significant improvement can be achieved by applying majority voting on multi-seasonal predictions resulted from multi-seasonal Sentinel-2 images, compared to the classification results averaged over single-seasonal predictions. Based

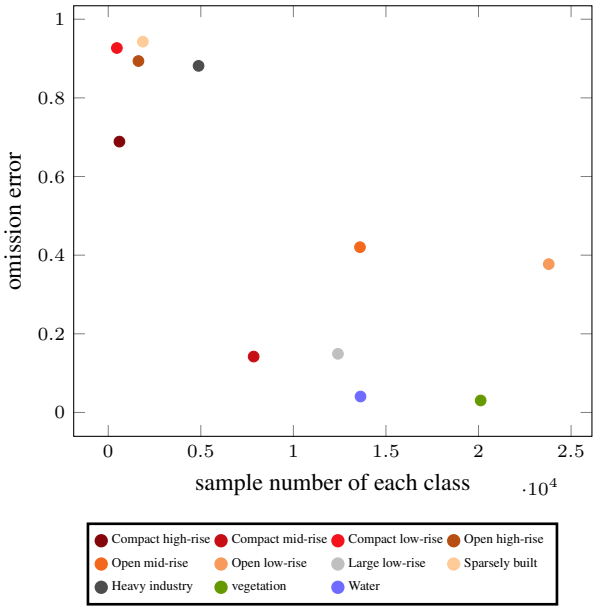


Figure 7. Relation between the class sample number and the omission error.

on the results, main challenges and possible solutions for large-scale urban LCZ classification are further discussed in detail. However, even majority voting does not provide the perfect solution to urban LCZ classification, which motivates investigations towards more advanced classifiers and the fusion of complementary data in the future.

ACKNOWLEDGEMENTS

This work is jointly supported by the China Scholarship Council, the Helmholtz Association under the framework of the Young Investigators Group SiPEO (VH-NG-1018, www.sipeo.bgu.tum.de), and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. ERC-2016-StG-714087, Acronym: *So2Sat*).

REFERENCES

Bechtel, B., Alexander, P. J., Böhner, J., Ching, J., Conrad, O., Feddema, J., Mills, G., See, L. and Stewart, I., 2015. Mapping local climate zones for a worldwide database of the form and function of cities. *ISPRS International Journal of Geo-Information* 4(1), pp. 199–219.

Bechtel, B., Demuzere, M., Sismanidis, P., Fenner, D., Brousse, O., Beck, C., Van Coillie, F., Conrad, O., Keramitsoglou, I., Middel, A. et al., 2017. Quality of crowdsourced data on urban morphologythe human influence experiment (huminex). *Urban Science* 1(2), pp. 15.

He, K., Zhang, X., Ren, S. and Sun, J., 2016. Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778.

Mou, L., Ghamisi, P. and Zhu, X. X., 2017. Unsupervised spectral-spatial feature learning via deep residual conv-deconv network for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*.

Qiu, C., Schmitt, M., Ghamisi, P. and Zhu, X. X., 2018. Effect of the training set configuration on sentinel-2-based urban local climate zone classification. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. in press.

Radoux, J., Chomé, G., Jacques, D. C., Waldner, F., Bellemans, N., Matton, N., Lamarche, C., D'Andrimont, R. and Defourny, P., 2016. Sentinel-2's potential for sub-pixel landscape feature detection. *Remote Sensing*.

Stewart, I. D., 2011. Local climate zones: Origins, development, and application to urban heat island studies. *Paper presented at the Annual Meeting of the American Association of Geographers*.

Stewart, I. D. and Oke, T. R., 2012. Local climate zones for urban temperature studies. *Bulletin of the American Meteorological Society* 93(12), pp. 1879–1900.

Tucker, C. J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote sensing of Environment* 8(2), pp. 127–150.

Xu, Y., Ma, F., Meng, D., Ren, C. and Leung, Y., 2017. A cotraining approach to the classification of local climate zones with multi-source data. In: *Geoscience and Remote Sensing Symposium (IGARSS), 2017 IEEE International*, IEEE, pp. 1209–1212.

Yokoya, N., Ghamisi, P. and Xia, J., 2017. Multimodal, multitemporal, and multisource global data fusion for local climate zones classification based on ensemble learning. *Proc. 37th annual symposium of the IEEE Geoscience and Remote Sensing Society (GRSS), Fort Worth, Texas, USA, 23-28 July.*

Zhu, X. X., 2018. So2sat LCZ42 dataset, to appear.

Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F. and Fraundorfer, F., 2017. Deep learning in remote sensing: a review. *arXiv preprint arXiv:1710.03959*.